Sherman Lam (sherman1)
Logan Ellis (logane)
Jack Buffington (jbuffing)
16-720 Computer Vision
October 12, 2016

3D Scene Reconstruction from Single Moving Camera

1 Project Overview

3D scene reconstruction is useful in robotics for understanding a robot's environment can aid in navigation through this environment. One typical approach is to use a combination of a GPS, laser scanners, and inertial measurement units to generate a map of the environment. However, this requires extensive, expensive, and specialized hardware. We would like to reconstruct a scene using video recorded from a standard monocular camera.

The goal of this project is to reconstruct a 3D scene using video recorded from a standard cell phone camera. This project will attempt to recreate the work of Newcombe and Davison ([8]). They present an algorithm that efficiently generates a dense model of a camera scene. It leverages a structure from motion algorithm called "Parallel Tracking and Mapping" (PTAM) to extract a high density of keypoints from each frame of the video feed. Their algorithm uses these points to first construct a low resolution model of the environment. Using groups of camera views with overlapping surface visibility, the model coarse model is refined to yield a dense depth map. Once this depth map is created for various scenes in the environment, the individual depth maps are stitched together to create a 3D map of the entire environment.

SFM is a relatively mature topic of study within computer vision, and, as such, there are a breadth of resources to draw from in pursuit of this project objective. Other papers that address similar problems and which we may refer to include [8], [1], [4], [9], [5], [6], [3], [7], and [2].

2 Technology

The authors of [8] produced their experimental data utilizing a camera recording at 640x480 resolution at 30Hz. The processing was performed in real-time on a desktop system utilizing a pair of discrete GPU's. This paper was originally published in 2010 and desktop computing performance has increased substantially since then. Ideally, a modern implementation could be performed in real-time on a single desktop GPU, with either resolution increased to 1280x720, or framerate increased to 60Hz. Video will be recorded utilizing a webcam connected to the processing computer to avoid bandwidth limitations that stem from telemetering of video data.

Algorithm development and implementation will be performed in MatLab, and supplemented by the accompanying computer vision toolbox only as necessary to maintain the appropriate scope of the project.

3 Timeline

The key deadlines for this project are:

- 11/9/16 - Finish extracting keypoints using PTAM and building low resolution model of the environment
- 11/13/16 - Check point for reviewing and understanding all underlying math.
- 11/28/16 - Finish basic, working construction of dense depth map.
- 12/3/16 - Finish algorithm optimization. Coding freeze.
- 12/5/16 - Final Presentation.

References

[1] Andrew J Davison. Real-time simultaneous localisation and mapping with a single camera. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1403–1410. IEEE, 2003.

[2] Andrew J Davison, Ian D Reid, Nicholas D Molton, and Olivier Stasse. Monoslam: Real-time single camera slam. *IEEE transactions on pattern analysis and machine intelligence*, 29(6):1052–1067, 2007.

[3] Andreas Geiger, Julius Ziegler, and Christoph Stiller. Stereoscan: Dense 3d reconstruction in real-time. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 963–968. IEEE, 2011.

[4] Georg Klein and David Murray. Parallel tracking and mapping for small ar workspaces. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 225–234. IEEE, 2007.

[5] Annika Kuhl, Christian Wöhler, Lars Krüger, Pablo dAngelo, and Horst-Michael Groß. Monocular 3d scene reconstruction at absolute scales by combination of geometric and real-aperture methods. In *Joint Pattern Recognition Symposium*, pages 607–616. Springer, 2006.

[6] Daniel Magree, John G Mooney, and Eric N Johnson. Monocular visual mapping for obstacle avoidance on uavs. *Journal of Intelligent & Robotic Systems*, 74(1-2):17–26, 2014.

[7] Philip F McLauchlan. A batch/recursive algorithm for 3d scene reconstruction. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 2, pages 738–743. IEEE, 2000.

[8] Richard A Newcombe and Andrew J Davison. Live dense reconstruction with a single moving camera. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1498–1505. IEEE, 2010.

[9] Marc Pollefeys, David Nistér, J-M Frahm, Amir Akbarzadeh, Philippos Mordohai, Brian Clipp, Chris Engels, David Gallup, S-J Kim, Paul Merrell, et al. Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision*, 78(2-3):143–167, 2008.